Also prove that

$$\rho^{-2} = 2W(\boldsymbol{\alpha}) = \|\mathbf{w}\|^2. \quad (7.69)$$

Note that for these relations to hold true, $\boldsymbol{\alpha}$ needs to be the solution of (7.29).

7.4 (Relationship Between $\|\mathbf{w}\|$ and the Geometrical Margin •) *(i) Consider a separating hyperplane in canonical form. Prove that the margin, measured perpendicularly to the hyperplane, equals $1/\|\mathbf{w}\|$, by considering two opposite points which precisely satisfy $|\langle \mathbf{w}, \mathbf{x}_i \rangle + b| = 1$.*

(ii) How does the corresponding statement look for the case of ν-SVC? Use the constraint (7.41), and assume that all slack variables are 0.

7.5 (Compression Bound for Large Margin Classification ◦◦◦) *Formalize the ideas stated in Section 7.2: Assuming that the data are separable and lie in a ball of radius R, how many bits are necessary to encode the labels of the data by encoding the parameters of a hyperplane? Formulate a generalization error bound in terms of the compression ratio by using the analysis of Vapnik [561, Section 4.6]. Compare the resulting bound with Theorem 7.3. Take into account the eigenvalues of the Gram matrix, using the ideas* from *[604] (cf. Section 12.4). Cf. von Luxburg et al., TR 101, MPI for Biological Cybernetics.*

7.6 (Positive Definiteness of the SVC Hessian •) *From Definition 2.4, prove that the matrix $Q_{ij} := (y_i y_j k(x_i, x_j))_{ij}$ is positive definite.*

7.7 (Geometric Interpretation of Duality in SVC [38] ••) *Prove that the programming problem (7.10), (7.11) has the same solution as (7.22), provided the threshold b is adjusted such that the hyperplane bisects the shortest connection of the two convex hulls. Hint: Show that the latter is the dual of the former. Interpret the result geometrically.*

7.8 (Number of Points Required to Define a Hyperplane •) *From (7.22), argue that no matter what the dimensionality of the space, there can always be situations where two training points suffice to determine the optimal hyperplane.*

7.9 (In-Bound SVs in Soft Margin SVMs •) *Prove that in-bound SVs lie exactly on the margin. Hint: Use the KKT conditions, and proceed analogously to Section 7.3, where it was shown that in the hard margin case, all SVs lie exactly on the margin.*

Argue, moreover, that bound SVs can lie both on or in the margin, and that they will "usually" lie in the margin.

7.10 (Pattern-Dependent Regularization •) *Derive a version of the soft margin classification algorithm which uses different regularization constants C_i for each training example. Start from (7.35), replace the second term by $\frac{1}{m}\sum_{i=1}^m C_i \xi_i$, and derive the dual. Discuss both the mathematical form of the result, and possible applications (cf. [462]).*

7.11 (Uncertain Labels ••) *In this chapter, we have been concerned mainly with the case where the patterns are assigned to one of two classes, i.e., $y \in \{\pm 1\}$. Consider now the*